

Folha de Pagamento - SPOBRAS

```
library(tidyverse)
library(na.tools)
library(readxl)
library(stringr)
library(lubridate)
library(pdftools)
library(data.table)
```

```
spobras<-fread("http://dados.prefeitura.sp.gov.br/dataset/3edd72cc-2f79-4c95-88e7-55744881b0a6/resource/
```

Folha de pagamento SPOBRAS - 2019

```
spobras<-as_tibble(spobras)
spobras
```

```
## # A tibble: 2,362 x 15
##   EMPRESA MES      ANO NOME  CARGO LOTAÇÃO ADMISSÃO NASCIMENTO VENCIMENTOS
##   <chr>    <chr> <int> <chr> <chr> <chr>    <chr>          <int> <chr>
## 1 SPObras Jane~ 2019 WALT~ ANAL~ NUCLEO~ 19/08/1~      1960 13482,77
## 2 SPObras Jane~ 2019 LUIZ~ AGEN~ NUCLEO~ 12/10/1~      1962 0
## 3 SPObras Jane~ 2019 VALD~ AGEN~ NUCLEO~ 10/04/1~      1965 6622,49
## 4 SPObras Jane~ 2019 MARC~ ANAL~ NUCLEO~ 14/07/1~      1959 23401,38
## 5 SPObras Jane~ 2019 MARC~ AGEN~ NUCLEO~ 03/08/1~      1956 6339,57
## 6 SPObras Jane~ 2019 JOAO~ AGEN~ NUCLEO~ 17/09/1~      1959 5255,93
## 7 SPObras Jane~ 2019 ELIA~ SECR~ GESTAO~ 01/12/1~      1955 9797,86
## 8 SPObras Jane~ 2019 SILV~ ANAL~ NUCLEO~ 28/12/1~      1962 14174,82
## 9 SPObras Jane~ 2019 MARC~ ANAL~ NUCLEO~ 01/11/1~      1958 14174,82
## 10 SPObras Jane~ 2019 ALFR~ AGEN~ NUCLEO~ 18/11/1~      1964 13508,67
## # ... with 2,352 more rows, and 6 more variables: ENCARGOS <chr>,
## #   BENEFÍCIOS <chr>, 'OUTRAS REMUNERAÇÕES' <int>, VÍNCULO <chr>, 'DETALHE
## #   VÍNCULO' <chr>, LIMINAR <chr>
```

CONTAGEM DE VALORES FALTANTES EM NOSSO DATASET.

```
sum(is.na(spobras))
```

```
## [1] 0
```

EXCLUINDO A VARIÁVEL LIMINAR

```
spobras<-spobras[,-15]
```

QUANTIDADE DE SERVIDORES EM NOSSO DATASET SPOBRAS

```
spobras %>% distinct(NOME) # RETORNA 241 NOMES
```

```
## # A tibble: 241 x 1
##   NOME
##   <chr>
## 1 WALTER MARTINS DE SOUZA
## 2 LUIZ CARLOS DE SOUZA BARBOZA
## 3 VALDELI DA SILVA
## 4 MARCIA APARECIDA TEIXEIRA DE CAMPOS CARVALHO
## 5 MARCOS SILVA ROCHA
## 6 JOAO PAULO RESPLANDES DE OLIVEIRA
## 7 ELIANA MARIANO FRANCO
## 8 SILVANA ISGROI CARVALHO
## 9 MARCIA HELENA NUCCI EIRAS GARCIA
## 10 ALFREDO LENZ
## # ... with 231 more rows
```

QUANTIDADE DE CARGOS DIFERENTES NO DATASET

```
spobras %>% distinct(CARGO) # RETORNA 44 CARGOS DIFERENTES
```

```
## # A tibble: 44 x 1
##   CARGO
##   <chr>
## 1 ANALISTA DE PROJETOS, OBRAS E SERVIÇOS-ENGENHEIRO
## 2 AGENTE DE PROJETOS, OBRAS E SERVIÇOS
## 3 AGENTE DE ADMINISTRAÇÃO
## 4 ANALISTA DE ADMINISTRAÇÃO
## 5 SECRETARIA
## 6 ASSISTENTE DE PROJETOS OBRAS E SERVIÇOS
## 7 ANALISTA DE AGRIMENSURA E CONSTRUÇÃO CIVIL
## 8 AUXILIAR DE ADMINISTRAÇÃO
## 9 ANALISTA DE PROJETOS OBRAS E SERVIÇOS - ARQUITETO
## 10 ADVOGADO
## # ... with 34 more rows
```

IDENTIFICANDO A QUANTIDADE DE VINCULOS NO DATASET

```
spobras %>% group_by('DETALHE VÍNCULO') %>% count('DETALHE VÍNCULO')
```

```
## # A tibble: 4 x 2
## # Groups:   DETALHE VÍNCULO [4]
##   'DETALHE VÍNCULO'      n
##   <chr>                <int>
## 1 Concursado NÃO ocupante de cargo em comissão    1371
## 2 Concursado ocupante de cargo em comissão         43
## 3 Não concursado NÃO ocupante de cargo em comissão  444
## 4 Não concursado ocupante de cargo em comissão    504
```

TRANSFORMANDO VARIÁVEIS VENCIMENTOS, ENCARGOS E BENEFÍCIOS EM VARIÁVEIS NUMÉRICAS

```
### VARIÁVEIS: VENCIMENTOS, ENCARGOS E BENEFÍCIOS ESTÃO COMO strings.. É PRECISO CONVERTE-LAS A NUMEROS
## (substituir a vírgula por ponto)
spobras$VENCIMENTOS<-str_replace_all(spobras$VENCIMENTOS,pattern = "[,]",replacement = ".")
spobras$ENCARGOS<-str_replace_all(spobras$ENCARGOS,pattern = "[,]",replacement = ".")
spobras$BENEFÍCIOS<- str_replace_all(spobras$BENEFÍCIOS,pattern = "[,]",replacement = ".")

### agora transforma-las em numeros com casas decimais
spobras$VENCIMENTOS<-as.double(spobras$VENCIMENTOS)
spobras$ENCARGOS<-as.double(spobras$ENCARGOS)
spobras$BENEFÍCIOS<-as.double(spobras$BENEFÍCIOS)
```

SEPARANDO DATASET POR TIPO DE VÍNCULO EMPREGATÍCIO

```
cncm<-spobras %>% select(NOME,CARGO,VENCIMENTOS,'DETALHE VÍNCULO') %>%
  filter('DETALHE VÍNCULO' == "Concursado NÃO ocupante de cargo em comissão") ## 1371 OBSERVAÇÕES

ccc<-spobras %>% select(NOME,CARGO,VENCIMENTOS,'DETALHE VÍNCULO') %>%
  filter('DETALHE VÍNCULO' == "Concursado ocupante de cargo em comissão") ## 43 OBSERVAÇÕES

ncnc<-spobras %>% select(NOME,CARGO,VENCIMENTOS,'DETALHE VÍNCULO') %>%
  filter('DETALHE VÍNCULO' == "Não concursado NÃO ocupante de cargo em comissão") # 504 OBSERVAÇÕES

nccc<-spobras %>% select(NOME,CARGO,VENCIMENTOS,'DETALHE VÍNCULO') %>%
  filter('DETALHE VÍNCULO' == "Não concursado ocupante de cargo em comissão") ## 444 OBSERVAÇÕES
```

DISTRIBUIÇÃO DE SERVIDORES POR VÍNCULOS E SALÁRIOS

```
# PROPORÇÃO DE SERVIDORES POR VÍNCULO
proporcao_servidores<-data.frame(Vinculo=c("Concursado Não Ocupante CM","Concursado Ocupante CM","Não C
      "Não Concursado Ocupante CM", "Total"),
      Qtd_servidores=rbind(nrow(cncm),nrow(ccc),nrow(ncnc),nrow(nccc),nrow(spobras)),
      Proporcao=rbind((nrow(cncm)/nrow(spobras)*100),(nrow(ccc)/nrow(spobras)*100),
        (nrow(ncnc)/nrow(spobras)*100),(nrow(nccc)/nrow(spobras)*100),(nrow(spobras)/nrow(spobras)*100))

### PROPORÇÃO SALÁRIO BRUTO POR VÍNCULO
proporcao_salarial_X_vinculo<-data.frame(Vinculo=c("Concursado Não Ocupante CM","Concursado Ocupante CM",
      "Não Concursado Ocupante CM", "Total"),
      Salario_Bruto_total=rbind(sum(cncm$VENCIMENTOS),sum(ccc$VENCIMENTOS),sum(ncnc$VENCIMENTOS),sum(nccc$VENCIMENTOS),sum(spobras$VENCIMENTOS)),
      Proporcao_salario_Bruto=rbind(sum(cncm$VENCIMENTOS)/sum(spobras$VENCIMENTOS),sum(ccc$VENCIMENTOS)/sum(spobras$VENCIMENTOS),sum(ncnc$VENCIMENTOS)/sum(spobras$VENCIMENTOS),sum(nccc$VENCIMENTOS)/sum(spobras$VENCIMENTOS),sum(spobras$VENCIMENTOS)/sum(spobras$VENCIMENTOS)))
```

TRANSFORMANDO EM TIBBLES

```
proporcao_servidores<-as_tibble(proporcao_servidores)
proporcao_salarial_X_vinculo<-as_tibble(proporcao_salarial_X_vinculo)
```

TABELA E GRAFICO: PROPORÇÃO SERVIDORES POR VINCULO

```
proporcao_servidores
```

```
## # A tibble: 5 x 3
##   Vinculo                Qtd_servidores Proporcao
##   <chr>                  <int>      <dbl>
## 1 Concursado Não Ocupante CM          1371      58.0
## 2 Concursado Ocupante CM              43       1.82
## 3 Não Concursado Não Ocupante CM      444      18.8
## 4 Não Concursado Ocupante CM         504      21.3
## 5 Total                          2362     100
```

```
proporcao_servidores %>%
  ggplot(mapping = aes(x=Proporcao,y=Vinculo,fill=Vinculo))+
    geom_col(show.legend = FALSE)+
    ggtitle(label = "Proporção Servidores por Vínculo Empregatício")+
    xlab(label = "Porcentagem (%)")+
    ylab(label = "Vínculo Empregatício")
```

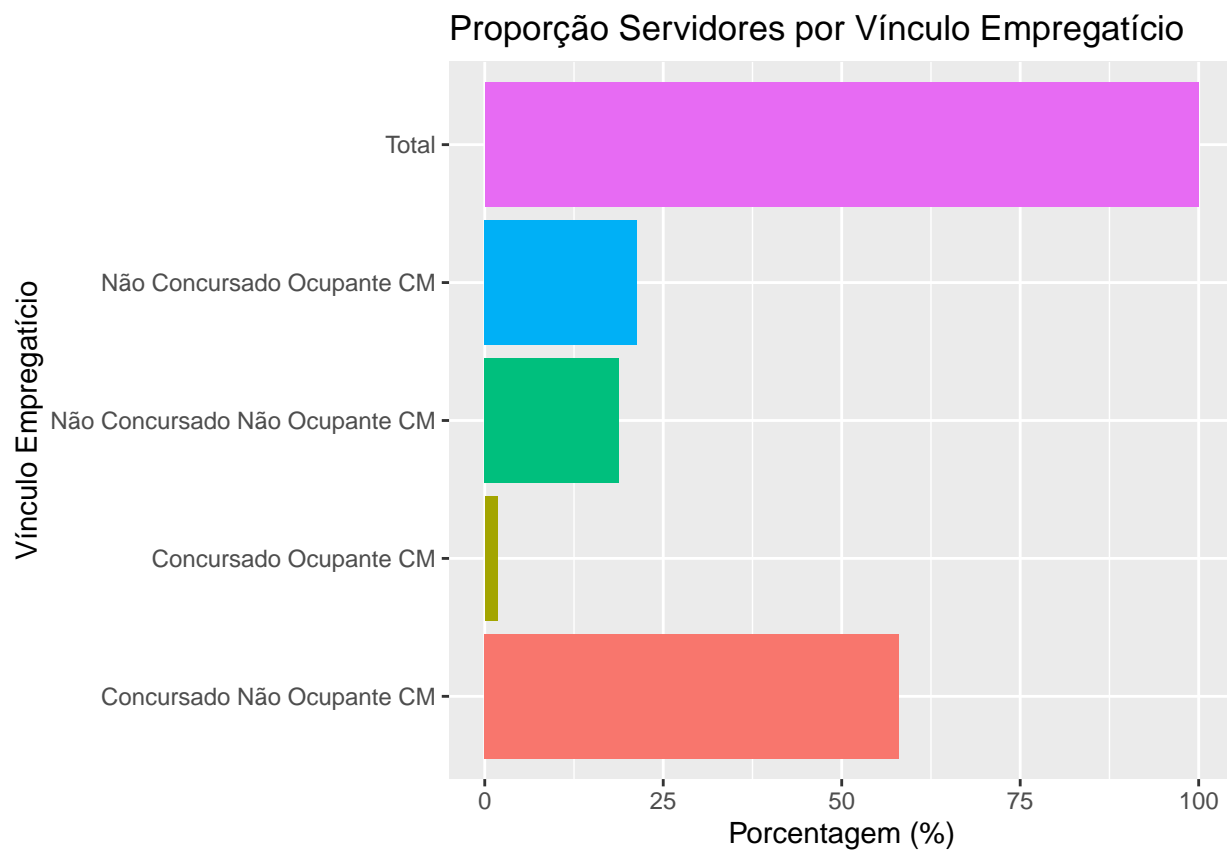


TABELA E GRAFICO: PROPORÇÃO SERVIDORES SALARIO BRUTO POR VINCULO

```
proporcao_salarial_X_vinculo
```

```
## # A tibble: 5 x 3
##   Vinculo                               Salario_Bruto_total Proprcao_salario_Bruto
##   <chr>                                <dbl>                <dbl>
## 1 Concursado Não Ocupante CM          11042321.             50.2
## 2 Concursado Ocupante CM              768121.              3.49
## 3 Não Concursado Não Ocupante CM     1416981.             6.44
## 4 Não Concursado Ocupante CM         8779982.            39.9
## 5 Total                             22007405.           100
```

```
proporcao_salarial_X_vinculo %>%
  ggplot(mapping = aes(x=Proprcao_salario_Bruto,y=Vinculo,fill=Vinculo))+
  geom_col(show.legend = FALSE)+
  ggtitle(label = "Proporção salário bruto recebido")+
  xlab(label = "Porcentagem (%)")+
  ylab(label = "Vínculo Empregatício" )
```

